

1. Big Data and the ethics of detail: the role of ethics work in the making of a cross-national research infrastructure for genetic research

Klaus Hoeyer, Aaro Tupasela, and Malene Bøgehus Rasmussen¹

1 INTRODUCTION

This chapter explores the mobilization of resources, technologies, patients, human capital, and biomaterials for international collaboration regarding genetic research. Our question is simple: What makes genetic material and health data flow, and which hopes and concerns travel along with them? What cannot travel, and what travels in unintentional ways? By focusing on the flows of material we elucidate the moral and social work that goes into the exchange of research materials and illustrate the divergences between this kind of work and the official ethics policies and frameworks that are supposed to guide it.

Our research site is a Danish laboratory that has longstanding experience of genetic research with international partners. We focus on two types of international collaboration. The first is a long-term collaboration with a research center in Pakistan aimed at studying rare autosomal recessive diseases. These diseases are more common in contexts with a tradition of consanguineous marriages. The manifest genetic disorders provide researchers with an opportunity to identify new disease-causing genes of wider relevance for understanding human biology. The second type of collaboration is an initiative taken by the Danish laboratory to establish a research consortium called The

¹ This chapter is a reworked version of a paper published in 2017 as “Ethics Policies and Ethics Work in Cross-national Genetic Research and Data Sharing: Flows, Nonflows, and Overflows” (2017) 42(3) *Science, Technology and Human Values*, 381–404. We thank Sage for allowing this reproduction.

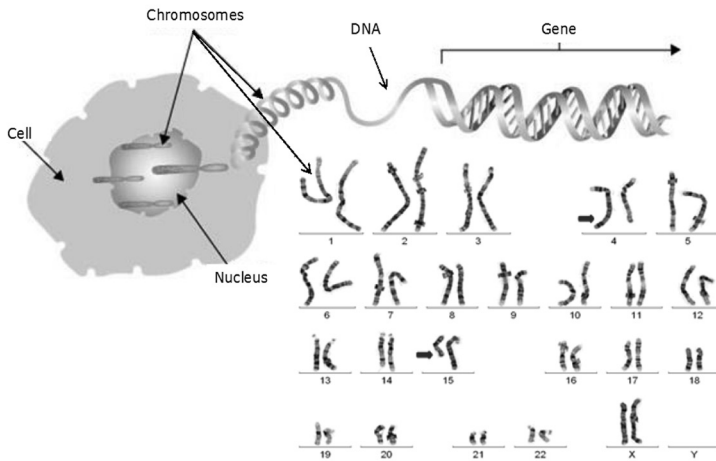
International Breakpoint Mapping Consortium (IBMC). The IBMC seeks to create a saturated map of balanced chromosomal rearrangements as a way to gain functional knowledge of the human genome. Most of these chromosomal rearrangements are rare; some are unbalanced, with missing or extra genetic material causing chromosomal disorders; others are so-called balanced rearrangements in which parts of chromosomes are moved or inverted, but in a way that does not necessarily cause any pathology (see Figure 1.1). Still, the chromosomal rearrangements may cause disease if the associated breakpoints truncate a gene or an area regulating gene expression. The goal is to compile a library of these chromosomal breakpoints from which new aspects of human genetics can be investigated. The library will serve as a biomedical platform,² allowing easier exploration of the role of rare chromosomal rearrangements in human disease. Diagnostic and research labs from more than 50 countries on six continents are participating in the endeavor by sending samples to the laboratory in Copenhagen where they are sequenced and added to the library. The lab draws upon its long experience with international collaborations when initiating new partnerships for IBMC.

Recent years have seen great emphasis on promoting international collaborations in genetic research. Policies are developed to promote data sharing, harmonization of international rules, and open access, and global alliances are formed.³ Mostly this work focuses on removal of so-called barriers to access, for example, by ensuring comparable demands of informed consent in different jurisdictions,⁴

² Peter Keating and Alberto Cambrosio, *Biomedical Platforms: Realigning the Normal and the Pathological in Late-Twentieth-Century Medicine* (MIT Press 2003).

³ P. Arzberger, P. Schroeder, A. Beaulieu, G. Bowker, L. Laaksonen, D. Moorman, P. Uhler, and P. Wouters, "Promoting Access to Public Research Data for Scientific, Economic, and Social Development" (2004) 3 *Data Science Journal*, 135–52; F. Colledge, B. Elger, and H. Howard, "A Review of the Barriers to Sharing in Biobanking" (2013) 11 *Biopreservation and Biobanking* (ahead of print); E. Dove, A-M. Tassé, and B.M. Knoppers, "What Are Some of the ELSI Challenges of International Collaborations Involving Biobanks, Global Sample Collection, and Genomic Data Sharing and How Should They Be Addressed?" (2014) 12 *Biopreservation and Biobanking*, 363–4. See, for example, the work of international networks such as Public Population Project in Genomics and Society (P3G), GenomEUtwin, and PHOEBE, and policy papers from, for example, the OECD. Note also the decision reached by the EU on May 27, 2016, to work towards a paradigm of "Open Science."

⁴ Barbara J. Evans and Eric M. Meslin, "Encouraging Translational Research through Harmonization of FDA And Common Rule Informed Consent Requirements For Research With Banked Specimens" (2006) 27 *Journal of Legal Medicine*, 119–66.



Notes: Left: The nucleus of a human cell contains 46 chromosomes, each of which is made of one long DNA double helix that contain hundreds to thousands of genes. Right: Balanced chromosomal rearrangements can be identified by examining the chromosomes in a microscope. They contain the correct amount of DNA and are in most cases considered harmless for the carrier. The image shows a balanced translocation between chromosomes 4 and 15. The translocation breakpoints are illustrated with arrows. Karyotype: 46,XX,t(4;15)(q21.3;q13).

Images with permission from NHS National Genetics and Genomics Education Centre.

Figure 1.1 Chromosomal rearrangements

shared rules of feedback of incidental findings,⁵ or principles of data sharing.⁶ Simultaneously, another set of policies has focused on ethical issues relating to the protection of autonomy and privacy, as seen for example in the World Medical Association's declaration on health data and biobanking. Such policies seek to control and restrict data flows. The two sets of policymaking generally develop independently, and tensions between them must be handled on the spot by those subject to their sometimes incompatible rules.

⁵ Alessandro Blasimme et al, "Disclosing Results to Genomic Research Participants: Differences that Matter" (2012) 12 *The American Journal of Bioethics*, 20–2.

⁶ Deborah Mascalonzi et al, "International Charter of Principles for Sharing Bio-Specimens and Data" (2016) 24 *European Journal of Human Genetics*, 1–8.

Some scholars have pointed out how data-sharing policies have been accused of neglecting the tension between data sharing and privacy protection.⁷ Others have focused on how network structures involve challenges to governance,⁸ as well as to funding.⁹ Based on a study of networks that were claimed to be successful, Mayrhofer and Prainsack have argued that international rules are less important than informal networks in creating harmonized scientific standards and that the scientific collaboration is coproduced with ethics.¹⁰ Some studies indicate that it is often surprisingly difficult to work across borders unless you already know the local collaborators.¹¹ Other strains of scholarship have articulated critiques of the many attempts to stimulate international collaboration. For example, international tissue exchange has been criticized for enacting a form of biopiracy wherein resources in low-income countries are made available without a fair return and where “benefit sharing policies” do little to alleviate the problems.¹² It has been pointed out that exchanges generate shifts in valuation of the tissue in the course of passing from research participants to local researchers and further on to international research partners,¹³ and such shifts have been problematized as involving “commodification”¹⁴ and exploitation.¹⁵

⁷ Jane Kaye, “The Tension between Data Sharing and the Protection of Privacy in Genomics Research” (2012) 13 *Annual Review of Genomics and Human Genetics*, 415–31.

⁸ I. Meijer, J. Molas-Gallart, and P. Mattsson, “Networked Research Infrastructures and Their Governance: The Case of Biobanking” (2012) 39 *Science and Public Policy*, 491–9.

⁹ R. Jean Cadigan et al, “Neglected Ethical Issues in Biobank Management: Results from a U.S. Study” (2013) 9 *Life Sciences, Society and Policy*, 1–13.

¹⁰ Michaela T. Mayrhofer and Barbara Prainsack, “Being a Member of the Club: The Transnational (Self-) Governance of Networks of Biobanks” (2009) 12 *International Journal of Risk Assessment and Management*, 64–81.

¹¹ Roger Bjugn and others, “What Are Some of the ELSI Challenges of International Collaborations Involving Biobanks, Global Sample Collection, and Genomic Data Sharing and How Should They Be Addressed?” (2015) 13 *Biopreservation and Biobanking*, 70–1.

¹² Cori Hayden, “Taking as Giving: Bioscience, Exchange, and the Politics of Benefit-Sharing” (2007) 37 *Social Studies of Science*, 729–58.

¹³ Warwick Anderson, *The Collectors of Lost Souls: Turning Kuru Scientists into Whitemen* (Johns Hopkins University Press 2008).

¹⁴ Hilary Rose, “From Hype to Mothballs in Four Years: Troubles in the Development of Large-Scale DNA Biobanks in Europe” (2006) 9 *Public Health Genomics*, 184–9.

¹⁵ See discussion in Melinda Cooper and Cathy Waldby, *Clinical Labor: Tissue Donors and Research Subjects in the Global Bioeconomy* (Duke University Press 2014); Hilary Cunningham, “Colonial Encounters in Postcolonial Contexts” (1998) 18 *Critique of Anthropology*, 205–33; Emma Kowal, Joanna Radin, and Jenny Reardon,

In this chapter, we are interested in the work that goes into making international collaborations operational in practice and the flows of information and biomaterial which this work facilitates or obstructs. Our study of collaborations is itself the outcome of collaboration between geneticists and social scientists and an instance of data sharing. After a note on methods, we describe “infrastructures for flows” and “ethics work” to clarify how we analyze genetic research collaborations. We then examine the practical everyday type of ethics work that makes international collaborations—and thereby Big Data—possible. More specifically, we wish to explore the everyday ethics work facilitating exchange by describing what is made to flow, the nonflows, and the overflows of communication and samples.

2 METHODOLOGICAL REFLECTIONS: HOW TO EXPLORE FLOWS?

Where should you go to explore the internal workings of collaborations among research participants, researchers, and their research partners?¹⁶ This collaboration between a geneticist (Malene Bøgehus Rasmussen) and two social scientists (Klaus Hoeyer and Aaro Tupasela) is the product of an interdisciplinary research program funded by the University of Copenhagen intended to investigate legal, ethical, and scientific challenges in crossborder sharing of biological material. The empirical data for our analysis are drawn from two interrelated research projects. The first is a genetic research project initiated by Professor Niels Tommerup to reexamine carriers of balanced chromosomal rearrangements conducted by Rasmussen (henceforth called the genetic study); the second is a social scientific study conducted by Hoeyer and Tupasela focusing on research participants, researchers, and international research collaborators affiliated with the genetic lab (henceforth called the social scientific study).¹⁷ The social scientific study consisted of observation

“Indigenous Body Parts, Mutating Temporalities, and the Half-Lives of Postcolonial Technoscience” (2013) 43 *Social Studies of Science*, 465–83; Jonathan Marks, “‘We’re Going to Tell These People Who They Really Are’: Science and Relatedness,” in Sarah Franklin and Susan McKinnon (eds), *Relative Values: Reconfiguring Kinship Studies* (Duke University Press 2001), 355–83; Jenny Reardon, *Race to the Finish: Identity and Governance in an Age of Genomics* (Princeton University Press 2005).

¹⁶ Bart Penders et al., “When Scientists, Scholars, Clinicians, Physicians and Patients Meet,” in Bart Penders et al., *Collaboration across Health Research and Medical Care* (Ashgate Publishing Ltd 2015), 312.

¹⁷ Studies of interdisciplinary collaboration suggest that it takes work and time to establish collaborations across the disciplinary boundaries of laboratory science and social science: see M. Albert, S. Laberge, and B. Hodges, “Who Wants to Collaborate with Social Scientists? Biomedical and Clinical Scientists’ Perceptions of Social

of, and semistructured interviews with, researchers in the lab in Copenhagen (Tupasela), as well as semistructured interviews conducted during 2014 in Denmark (Hoeyer) and Finland (Tupasela) with Danish research participants (Hoeyer) and Pakistani (Hoeyer) and Finnish (Tupasela) collaborators. The interview framework explored the hopes and concerns that research participants, researchers, and research collaborators may or may not attach to the use of the samples and related information as they move across the globe. Quotes are translated from Danish and Finnish by the authors.¹⁸

Research participants' views are primarily explored through interviews with Danish individuals participating in the genetic study. The interviewees were recruited through the genetic project and the interviews were carried out by Hoeyer.¹⁹ Adult carriers of balanced chromosomal rearrangements who lived in Denmark were invited to participate in the genetic study, which included a questionnaire on health-related issues. One question asked participants if they were interested in being contacted about participating in the social scientific study to offer their opinions about the research. Research participants for the social scientific study were chosen by the authors to maximize variation along classic demographic variables such as age, occupation, education level, and place of residence, as well as differences in response to the invitation to participate in the original genetic study. Enrolment of participants in the genetic and the social scientific projects followed two separate research protocols approved by the Danish Data Protection Agency.²⁰ As part of the genetic study, Rasmussen had been in contact with several research participants and knew that some participants had responded with enthusiasm, some did not seem to pay much attention to it, and others had expressed anxiety and concern. Such differences in attitude toward the genetic study also were

Science," in Bart Penders et al, *Collaboration across Health Research and Medical Care* (Ashgate Publishing Limited 2015), 59–80. We, too, had to gradually negotiate the terms of our own collaboration and establish relationships to make meaningful the sharing of facilities and research network in Denmark, Finland, and Pakistan. In this way, our own collaboration came to mirror the phenomenon that the social scientific study had set out to describe, namely how flows of data are established through social work, though we only gradually realized this point in the course of writing the chapter.

¹⁸ We would like to thank Zainab Sheikh for helping with the Urdu translation.

¹⁹ The genetic study was approved by the Danish Data Protection Agency (J.nr. 2012-54-0053) and by the Danish regional Research Ethics Committees (J.nr. H-KF-2006-5901). Health-related information obtained through the genetic study was never disclosed to or discussed within the forum of the social scientific study, in agreement with the Danish Data Protection protocol (J.nr. 2012-54-0053).

²⁰ J.nr. 2014-41-3055: Global Genes, Local Concerns: Legal, Ethical and Scientific Challenges in Cross-National Biobanking and Translational Exploitation.

used to select potential candidates for semi-structured interviews.²¹ A total of 32 participants from the genetic study were invited to engage in the social scientific study. Two opted out, one did not reply, and six engaged only in email-based discussions, while 23 were interviewed either in their homes, in the office of Hoeyer, or over the phone, depending on their personal preference and logistical opportunities. The interviews lasted between 15 minutes and almost 3 hours. Most took about an hour. After approximately 15 interviews, we began reaching data saturation in the sense that we heard similar hopes and concerns repeated. All interviews were recorded and transcribed.

Interview candidates among the Danish researchers and their collaborators were identified using the snowball method, as well as through the identification of key partners in the IBMC. Beginning with the head of the genetic lab and a key staff member, we moved on to research partners in Finland and Pakistan, and some of their staff members. The Pakistani researchers either worked in Copenhagen or came on regular visits, while the Finnish partners were interviewed in Finland. We were only able to do interviews with Danish research participants, because we were able to ask for their permission through the original enrolment questionnaire used for the genetic study. The rules set by the Danish Data Protection Agency and Research Ethics Committee system have shaped what we could do with this analysis, and so understanding the rules became an interesting parallel study that also informed the analysis. While we have collaborated to understand how we might enhance the ethics of research collaborations, the ethics policies that are expected to guide collaborations have often felt like arbitrary restrictions. Rather than stimulating consideration of how to respect the autonomy and integrity of the interviewees, we have often found ourselves speaking about what we were *allowed* to do. Ethics rules, by way of being “rules,” have thus stimulated conversations about legality rather than values; about limits to what we could do, rather than thoughts about what we should do. We have used this observation as an impetus to reflect on the ethos of our own work and how it relates to the rules that are supposed to guide it and ensure its legitimacy.

²¹ To ensure compliance with rules set by the Data Protection Agency, neither Hoeyer nor Tupasela was informed about who fit which criteria. Similarly, the anonymity of the interviewee was upheld the other way round. Interviewees were informed that the researchers from the genetic study would not know who said what to Hoeyer and that their participation in the genetic study would not be affected by their decision to give interviews. As a member of the genetic study, Rasmussen has not seen the individual transcripts and was only informed about interview contents in anonymized form.

3 DEFINING ETHICS WORK AND INFRASTRUCTURAL FLOWS AS ANALYTICAL OBJECTS

By focusing on flows we are inspired more broadly by what Sheller and Urry have described as a “mobility turn” in the social sciences.²² The mobility turn has involved paying increased attention to what moves, how it moves, and what are seen as legitimate and illegitimate movements. Flows of samples and information are coproduced with an infrastructure facilitating the movement. Following Star and Ruhleder’s seminal paper on research infrastructures,²³ we think of infrastructures as activities rather than things. Everyone and everything is simultaneously engaged in many different relations, hence there is not *one* infrastructure, but multiple interrelated infrastructures enacted through different practices. The Danish laboratory of course still depends on a material infrastructure of freezers and information software, but such hardware is simultaneously engaged in many other flows.

A flow of material depends on much more than freezers.²⁴ Groundwork is required for people to donate samples and for researchers to ship the samples to Copenhagen. We think of this work, aimed at enacting sustainable relations, as “ethics work.” As argued by Mayrhofer and Prainsack,²⁵ the scientific collaboration is coproduced with ethics through informal networks. Still, the practical and mundane ethics work that is necessary to make material and information flow continues to receive limited attention. With “ethics work” we build on the social study of medical ethics as a practical activity, not just abstract values or principles.²⁶ This involves focusing on what people do, rather than what they think they ought to do. In the following, we wish to focus on the otherwise tacit aspects of this work, because we believe it is central to

²² Mimi Sheller and John Urry, “The New Mobilities Paradigm” (2006) 38 *Environment and Planning A*, 207–26; John Urry, *Sociology beyond Societies: Mobilities for the Twenty-First Century* (Routledge 2000).

²³ Susan Leigh Star and Karen Ruhleder, “Steps toward an Ecology of Infrastructure: Design and Access for Large Information Spaces” (1996) 7 *Information Systems Research*, 111–34.

²⁴ S Leonelli, “What Difference Does Quantity Make? On the Epistemology of Big Data in Biology” (2014) 1 *Big Data & Society*, 1–11.

²⁵ See n 9.

²⁶ Charles L. Bosk, *Forgive and Remember: Managing Medical Failure* (University of Chicago Press 1979); Daniel F. Chambliss, *Beyond Caring: Hospitals, Nurses and the Social Organization of Ethics* (Chicago University Press 1996); P. Wenzel Geissler et al, “‘He Is Now Like a Brother, I Can Even Give Him Some Blood’—Relational Ethics and Material Exchanges in a Malaria Vaccine ‘Trial Community’ in The Gambia” (2008) 67 *Social Science & Medicine*, 696–707.

international research collaborations. Whereas Sleeboom-Faulkner and Patra emphasize the exploitive potential of international collaborations that bridge and even thrive upon inequalities,²⁷ we focus here on what people do to engage the exchanges. We wish to elucidate the tacit work done to respect the involved stakeholders, because we believe it should be acknowledged as central to good and sustainable research ethical practices.

How should we conceptualize the objects around which collaborations revolve? Are we dealing with “a biobank,” for example? The notion of a biobank is problematic because samples used in international collaborations often belong to many different “banks.” Samples derive from numerous different sources, such as hospitals and research projects, and they are used for very different purposes. They form part of very diverse flows. In some cases the samples are not even stored at the lab: they are completely used up in the course of the research, leaving only data behind. The concept of “biobank” might be both too singular (one set of samples in one place), too static to capture the sense of flow (it indicates accumulation), and too informed by one type of purpose (research) to capture all the involved flows and uses. Would the concept of platform then work as an alternative to biobank? Indeed, Keating and Cambrosio suggested the concept of the biomedical platform to tie in the social and material aspects of networks in ways that simultaneously captured the agential capabilities of such networks.²⁸ However, if you ask around in the Copenhagen lab, the word platform figures in a quite specific sense: it signals, for example, the equipment used for sequencing (for example, a technician explained: “The Illumina platform can sequence 32 mate-pair samples in one 11-day run”). Other technical “platforms” are used for other purposes, such as “genome browsers” to look up what is already known about a particular genetic sequence. The IBMC map is expected to serve a similar function. The concept of platform therefore already holds a slightly different meaning, and it poorly captures the range of social activities and all the tacit work through which flows are enacted in international collaborations. We therefore suggest focusing on what collaborations aim to do—enact *flows* of biological material and data—and the work that goes into making it happen. With a focus on flows we deliberately bring back in the human actors, and the stories that matter to them.

We begin by describing the practical work going into the creation of a flow of genetic material and probe the hopes and concerns that travel along with

²⁷ Margaret Sleeboom-Faulkner and Prasanna Kumar Patra, “Experimental Stem Cell Therapy: Biohierarchies and Bionetworking in Japan and India” (2011) 41 *Social Studies of Science*, 645–66.

²⁸ See n 1.

it. To do so, we first reflect on the different strategies for identification of research participants and then provide examples of how participants are enrolled in Denmark, Finland, and Pakistan.

4 FLOWS: WHAT FLOWS AND WHAT MAKES IT FLOW?

The Danish laboratory recruits research participants in different ways depending on the options available in each country. In Denmark, potential genetic research participants (that is, carriers of chromosomal rearrangements) are identified through the Danish Cytogenetic Central Register.²⁹ In Finland there is a similar register, but it is run by a researcher and all access to the register goes through this person. In Pakistan, recruitment of research participants is based more on a snowball method and happenstance, since there are no centralized registers or databases for rare diseases. The Pakistani samples are not currently used for the IBMC map because these samples represent different kinds of genetic disorders. But the mode of collaboration has informed the work with the IBMC map, and the samples collected are used toward the same overall purpose of understanding genetic causes of disease.

To begin the genetic study, the lab in Denmark first seeks approval from a Research Ethics Committee and the Data Protection Agency. It then receives a list of carriers of chromosomal rearrangements from the Danish Cytogenetic Central Register. However, the lab does not approach the potential research participants right away. First, each diagnostic facility having entered patients into the register is approached and an agreement is made with them on how to approach potential research participants. They need to determine whether the carriers are aware of their carrier status, as for ethical reasons they do not wish to contact people who have not previously been informed of their status. This information is gathered from the original medical record describing the chromosome analysis performed—a point to which we return in the next section. The participants are invited to participate through a letter containing information describing the project, a consent form, and a health-related questionnaire. The participants are also informed that they might be invited to donate a biological sample such as a blood sample. The lab keeps records of all formal aspects of the research project as covered by the research protocols for the Research Ethics Committee and the Data Protection Agency, including consent status of participants, whether participants wish to be informed about clinically relevant results from the project, and if a biological sample has been

²⁹ Johannes Nielsen, *The Danish Cytogenetic Central Register: Organization and Results* (Thieme-Stratton 1980).

donated to the project. This all follows official guidelines and international standards. In addition, however, the lab finds it necessary to record questions or concerns relating to the project articulated by the research participant in order to ensure that more personal concerns are respected. It might be that a particular family member is not informed about his or her carrier status, and therefore should not be approached by the research team. In this way, some hopes and concerns travel along with samples and health data, though they flow through separate record systems.³⁰

What do the Danish research participants think about such records of social and moral concern? In our interviews, some participants expressed very strong concerns about particular family members who should not be informed about the identified chromosomal rearrangements. For them, keeping records of those concerns was essential to their trust in the researchers, though it was not conceptualized as a particular task demanding its own records. Researchers are just expected to *know*. One man had decided not to tell one of his children about their shared carrier status, and when he received the invitation to participate in research he was cited as thinking: “What then might happen tomorrow? ... How about my grandchildren; I don’t know, are they also gonna be summoned? Well, then I need to inform [the child] first ... there is an element of fear in all of this.” He relied on the researchers to realize that only some family members were aware of their carrier status. Interestingly, it was more common for the research participants to state that they were happy to participate in genetic research as long as researchers accessed *only* biomedically relevant information, and knew nothing about their personal hopes and concerns. One woman, for example, said that she did not want the researcher to know anything about her choices and preferences, but they could use her blood for whatever they wanted. She continued: “I think technologies have progressed very far, but they probably cannot figure out anything *personal* about me by looking at my blood ... So they can do whatever investigations that they want on my old blood. I don’t have any problem with that.” Ironically, for these research participants, keeping records of their stated hopes and concerns, which is meant to

³⁰ Terms of participation in biomedical research projects are covered by guidelines from the Research Ethics Committee system and the Danish Data Protection Agency. These agencies prescribe the formal requirements for informed consent, including information about research aims; benefits and risks of participation; terms of use of biological samples, including storage and limitations of use; confidentiality; and voluntariness. However, the Research Ethics Committee system and the Danish Data Protection Agency provide only limited instructions on how to address and keep track of more subtle issues such as the moral concerns uttered by the research participants. Recording what people say has become a tacit form of ethics work emerging out of practical experience and moral engagement with research participants over many years.

protect their privacy and respect their autonomy, is aligned with a precarious sense of concern, while genetic information and health-related data are seen as less problematic. A record of moral concerns and participant choices is nevertheless instrumental for respecting their position. Keeping records of personal choices is part of demarcating what should and should not flow, but these records themselves are not supposed to flow anywhere.

In Finland, the register for balanced chromosomal rearrangements is run and managed by one researcher who has visited all the hospital regions that have a genetic counseling and/or analysis unit. At these units, he has collected all the information available about individuals and families who have been identified with chromosomal rearrangements and entered them into a register, which his group uses for research. The register is not part of the national infrastructure but is run by this researcher. Once identified and entered into the register, the patients were first sent a letter regarding the research and given the possibility of contacting and discussing their problems with a genetic counselor or other senior physician. Later, if they consented to participate in the research, they were sent a sample collection kit that they could take to their local clinic to have a sample taken, which was then returned to the researcher. This collection kit included the contact details of a research group member who could be contacted if there were questions regarding the sampling procedure. During the process of collecting samples from about 100 patients who had requested the sample kit, one of the research group members noted that she was contacted by some patients who wanted to tell her their “stories” and experiences of living with their condition. In many cases, this related to having had a miscarriage, but other medical conditions were also discussed. Patients apparently wanted to share their experiences with the researcher in a way that was not sufficiently represented or covered by the sample collection or medical history approach. Though this information needs to flow for the patients to feel confident, it does not travel along with the samples to Copenhagen, as we shall discuss in the section on nonflows below. Some patients were also concerned about how information produced in the research would be shared, since they did not want information to be shared with family members. The flows were conditioned on nonflows. Other patients wanted additional information regarding chromosomal rearrangements and the research itself. In such cases, the flow of genetic material in one direction rested on genetic knowledge flowing in the opposite direction to the research participants. Again, social skills are needed to create genetic material flows: infrastructures depend on important, albeit tacit, forms of ethics work which do not feature in the official documents of collaboration.

The recruitment process in Pakistan is as dependent on social skills, or what we call “ethics work,” as it is in Denmark and Finland—but in significantly different ways. Samples are typically collected from people who do not have access to any of the genetic counseling services through which the Danish and

Finnish participants are identified. As mentioned previously, there is a high prevalence of genetic disorders, which is seen by scientists as related to high rates of intrafamily marriages. Scientists explain how, to gather samples and medical information, they first have to establish a relationship with these families. They do so through research students who are recruited to work for the institute, partly based on where they come from: by having students from a wide range of local areas, the institution has a network of people familiar with relevant local knowledge. Once contact is established through the local student, the institution can offer various forms of genetic counseling services, and in their interviews the scientists describe how they sometimes bring family members from rural areas to hospitals in major cities where they can undergo more sophisticated diagnostics and treatment offers.³¹ Diagnostic, treatment, and research activities are thus fully intertwined, though certain bioethical ideals tend to suggest that they are better kept safely apart.³² All these activities have to take place in order to create the relationships needed for information and samples to begin their travel.³³ When samples travel on to Copenhagen, it is often in the company of PhD students or postdoctoral researchers undertaking training in Denmark. These researchers bring with them an awareness of the stories of the people in whom the samples originated, and they are concerned about the ability to return information to the patients and their families.

In many ways, the samples traveling across the globe thus remain attached to the hopes and concerns through which they were first produced. However, some samples are sent without stewards, and samples are used for research that the Pakistani researchers are not themselves directly involved in executing. The international collaboration here relies on trust between the research partners (rather than between researcher and research participant). When characterizing this trust, one of the researchers used the word *Bharossa* [بھروسہ]: “It is, you could say, a super degree of trust, *Bharossa* [بھروسہ] ... When you rely on somebody that you have the highest level of trust in, then it is *Bharossa*!” *Bharossa* typically belongs to a religious idiom signaling faith, and the researcher thus emphasizes a very different kind of social contract than what can be contained in Material Transfer Agreements (MTAs) or

³¹ Prasanna Kumar Patra and Margaret Sleeboom-Faulkner, “Informed Consent and Benefit Sharing in Genetic Research and Biobanking in India: Some Common Impediments in Practice,” in Peter Dabrock et al (eds), *Trust in Biobanking* (Springer 2012), 237–56.

³² Paul S. Appelbaum et al, “False Hopes and Best Data: Consent to Research and the Therapeutic Misconception” (1987) 17 *The Hastings Center Report*, 20–4.

³³ See also Anderson, n 12; Emma Kowal, “Orphan DNA: Indigenous Samples, Ethical Biovalue and Postcolonial Science” (2013) 43 *Social Studies of Science*, 577–97.

policies of open access. Hoeyer asked the researcher what he would consider a *breach* of trust. He replied: “If they send the sample to a third party without my knowledge, which they have never done. Whenever they correspond, our collaborator, with a third, fourth, or even fifth research group, they always ask me.” Samples are made to travel in a sociomaterial infrastructure that involves much more than data storage and exchange. Notions of open access as a plain matter of sharing everything with anyone seem to ignore how collaborations are socially embedded. There is a morality built into the ties between people, and these ties restrict the flows. This takes us to the nonflows of research collaborations.

5 NONFLOWS: WHAT CANNOT FLOW AND WHICH FLOWS ARE STOPPED?

In addition to the formally regulated types of nonflows, such as ensuring confidentiality by not circulating patient names or social security numbers among researchers, there are other, more informal aspects of daily activities that stop material and information from flowing. In this section we look at nonflows, first by looking at how the creation of social ties between research partners can prevent them from entering other partnerships. We then examine examples of nonflows of information in the form of what can be termed “strategic ignorance,”³⁴ where some aspects of research collaboration are not revealed to research participants because they are expected to dislike them.

Concerning mutual obligations that imply nonflows, a Pakistani researcher noted that he was “morally obliged” to continue working with the Danish lab, and therefore unable to initiate collaborations with others when these would collide with the research interests of the Danish lab. Similarly, the diagnostic labs delivering samples for the IBMC map are expected not to support competing attempts of breakpoint mapping, although there are no formal rules against that. So the establishment of a flow in one direction creates interdependencies and limits to other flows. We might say that “trust” implies an outside border delineating what must be excluded from the relationship. It is well known that researchers often try to ensure a lead position by way of controlling resources, while both patient organizations,³⁵ on the one hand, and funding agencies and authorities (such as the Organisation for Economic Co-Operation and

³⁴ Linsey McGoey, “Strategic Unknowns: Towards a Sociology of Ignorance” (2012) 41 *Economy and Society*, 1–16.

³⁵ Georg Lauss, “Sharing Orphan Genes: Governing A European-Biobank-Network For The Rare Disease Community,” in Peter Dabrock et al (eds), *Trust in Biobanking* (Springer 2012), 219–35.

Development (OECD)),³⁶ on the other, seek to foster wider sharing of samples. It can be interpreted as a simple conflict of interest. However, what we suggest here is that some nonflows might not be about ensuring an edge in terms of competition at all. We might have to acknowledge that relationships of trust that are stable enough to facilitate flows of sensitive material cannot easily be made “open access.”

The notion of an outside border is also embedded in rules of *confidentiality* as they relate to research participants, irrespective of the national context in which the collection takes place. As a consequence, the records of personal histories and concerns that each lab keeps do not flow. When, for example, samples are sent from Finland, they are coded and contain no information on the person from whom they were collected. Even when Pakistani researchers travel to Copenhagen and work on samples they have collected themselves, they rarely share the personal histories of the participants with the lab technicians or other research partners. At some point, the samples travel unaccompanied by social hopes and concerns, partly because rules of confidentiality ensure that nothing personal is communicated. In this way, ethics rules can limit the continued awareness of local moral concerns. The lab in Copenhagen collaborates with another lab in which scientists do experiments on zebrafish to explore functional capacities of the identified chromosomal rearrangements. Illustrating how the samples become stripped of individual histories, a staff member in this biology lab stated in a conversation with us that he never considered the people behind the samples. Thus, the collaborative research is produced through social skills, as illustrated previously, but it simultaneously disentangles genetic information from social histories.

Finally, there are elements of the sample collection that involve what is described in the literature as strategic ignorance: planned nonflows of information.³⁷ In order to operate in rural areas of Pakistan, researchers have to take into account how their actions can be interpreted and thus avoid sharing particular forms of information. Following the capture of Osama bin Laden, and public statements about his identification through the collection of blood samples

³⁶ Organisation for Economic Co-Operation and Development (OECD) “OECD Principles and Guidelines for Access to Research Data from Public Funding” (2007), www.oecd.org/science/sci-tech/38500813.pdf (accessed 25 June 2018).

³⁷ Paul Wenzel Geissler, “Public Secrets in Public Health: Knowing Not to Know while Making Scientific Knowledge” (2013) 40 *American Ethnologist*, 13–34; Linsey McGoe, “Strategic Unknowns: Towards a Sociology of Ignorance” (2012) 41 *Economy and Society*, 1–16; Robert N. Proctor, “Agnotology: A Missing Term to Describe the Cultural Production of Ignorance (and Its Study),” in Robert N. Proctor and Londa Schiebinger (eds), *Agnotology: The Making and Unmaking of Ignorance* (Stanford University Press 2008), 1–36.

from his son during a polio vaccine campaign, the collection of samples by outside agencies in that area has been surrounded with anxiety, questions, and doubts. One of the researchers taking part in the collection explained: “If you go to that area, and any member of the team is not known by the people where we go, then they ask ‘Who are you, what is your objective, what will you do, will you give this blood to the CIA or what is the purpose of this?’” Collaboration with a Western lab does not help build trust in this environment. Furthermore, Denmark is discussed as a problematic country following the publication there of a series of cartoons depicting the prophet Muhammad. As a consequence, sometimes field staff decide to downplay the international aspects of their science collaborations when explaining their research to local participants, or they provide explanations only after a proper relationship has been established. Then the thought of samples traveling to Denmark can in fact be comforting, a Pakistani researcher explained, because it is taken as proof that the researchers take their condition seriously. Again, this initial nonflow is not meant to deprive the research participants of an “autonomous choice.” Rather, it is seen by the researchers as a way of expressing care for local sentiment in order to enroll them into research that can hopefully help them: giving them information about the condition and their potential carrier status can be essential for family planning. Information is not seen as a one-off thing that should be given prior to all involvement, as some consent policies suggest. Informational needs are seen as developing along with the relationship. The flows of information that eventually make up the infrastructure for the genetic research are intertwined with nonflows, protective boundaries, outside borders, and silences. Such nonflows are as socially embedded as fully articulated hopes and concerns.

The flows (and nonflows) we have described produce genetic knowledge that can operate free from the specific social histories of the research participants. For the most part, researchers and lab technicians working with samples in the lab consider their work routine and avoid attributing any personal attachment to them. However, it is important to note that there are several cases of overflows, where unintended information and unwarranted hopes and concerns travel along with the samples into the labs and from the labs into the homes of the people donating the samples. It is to such overflows that we now turn.

6 OVERFLOWS: WHAT TRAVELS UNINTENTIONALLY?

Though open access and data sharing policies intend to provide access only to scientifically relevant material, some forms of social meaning might also flow along with them. Several lab technicians working with diagnostic samples,

for example, remarked that it was difficult not to consider that a positive test result could have significant implications for the affected family. The fates of the unknown people taking the diagnostic tests thus come to matter even when the identities of the persons remain unknown. In such situations, genetic information overflows with social concerns.

More significant, perhaps, are the types of overflow that move in the other direction: from the research activities to the research participants. One of the Pakistani researchers described how collection of samples could interact with negotiations of etiology. When collecting samples, she found herself negotiating far more than the terms of participation in research: she would find herself introducing novel conceptions of the causes of illness that move the understanding of, for example, schizophrenia from a religious realm into a secular biological realm. The research activity thus leaves behind new narratives of illness in the involved families. In other cases, narratives of genetic illness can cause a fear of stigma: irrespective of the scientific explanations given, some families experience reactions from others who think that it is better to avoid marital relations with particular strains of the family tree. The global networks of science will always interact with these very personal stories in unwarranted ways. Science is unruly and not subject to total legal, ethical, or scientific control. Its material infrastructure overflows with meaning.

One might expect the mountainous regions of Pakistan to harness greater interpretive flexibility with respect to the scientific narratives than is the case among the literate Danish population. However, local interpretations are a consistent feature of the stories told also by Danish research participants. Several Danish research participants, for example, explained how the questionnaire they filled in as part of their participation in the genetic study led them to associate their chromosomal rearrangement with a range of personal characteristics. Some said they had come to feel more closely related to other family members with the same rearrangement. Others describe how they used the questionnaire as a source of information about “what science now knows,” and used it to understand the cause of particular diseases among family members, for example. The research activity thereby sets in motion narratives about health and illness that are founded not on science but on local interpretations by people looking for explanations. People use pieces of information picked up during the research activity to symbolically make sense of their world.³⁸

There are also people with balanced chromosomal rearrangements who wonder whether they can be organ or blood donors or whether their genetic

³⁸ Klaus Hoeyer, “Traveling Questions: Uncertainty and Nonknowledge as Vehicles of Translation in Genetic Research Participation” (2016) 35 *New Genetics and Society*, 351–71.

makeup might cause danger to potential recipients. One man, for example, had registered an optout for organ donation despite actually supporting organ donation, because he feared his rearrangement could cause an organ rejection. In the course of an interview he said: “And then some poor fellow, who is already very ill, gets my organ, ‘hurrah!’ And then two weeks later ‘sorry, mate, it didn’t work out, it was rejected’. I would feel like a villain [slyngelagtig].” Following the interview, Hoeyer informed him that he had little need to worry about such a rejection; after getting confirmation from Rasmussen based on an anonymized description of the case, Hoeyer was able to further console the man, who then wrote back that he would change his decision to allow organ donation. Again, such overflows emerge as people apply the information acquired through research participation to concerns of their own that are not covered by official guidelines. And again, the research participant was consoled thanks to our collaboration, and despite—rather than because of—the official guidelines, which were hindering passage of information from the social scientific study to the genetic study.

Researchers in the genetic study are keenly aware of potential overflows, but they are not in a position to control them. Importantly, the official ethics policies focus on what people *should* know *prior* to accepting the invitation to research, not on how to work with such unintended *consequences* of research participation. Such work demands funding and continued dedication. It is a very different kind of work than that involved in making data available in open access depositories, but if the moral concerns of donors count it cannot be neglected. Researchers, research participants, and funding agencies may operate with very different definitions of what constitutes appropriate sharing and what constitutes appropriate flows, preferred nonflows, and overflows. It largely remains a personal query for researchers to contemplate these issues and to figure out how to respond to them.

7 CONCLUSION

We have described how recent years have seen a substantial policy emphasis on stimulating international collaboration in genetic research through data sharing policies and global regulatory harmonization. We have argued that when policymakers think of the task as one of encouraging researchers to “share” and “remove barriers,” they might misconstrue research collaboration as a simple matter of providing access to preexisting freezers and databases. Data sharing does not materialize without work, however, and this work is socially embedded. To promote data sharing, it is important to understand the social mechanisms involved. Networks that support the flow of sensitive material and information may not easily enable “open access” because the social mechanisms that facilitate the flow depend on clear expectations about mutual

social obligations. They build on relationships, not just infrastructures of data availability. Big Data depends on an ethics of detail. Though we have explored genetic research collaborations, similar social mechanisms also might shape research in other fields. Our findings suggest that effective social relations constitute an important prerequisite for data sharing. The dynamics of flows, nonflows, and overflows are configured in different ways, depending on the type of sharing network that is in question. Where policies of “open science”³⁹ tend to imagine that material and information will flow if only an infrastructure is provided (and a demand for sharing is installed), we instead suggest paying increased attention to the work going into the making of flows.

Furthermore, resistance to data sharing can reflect concerns other than self-interested attempts to ensure a competitive edge. We have argued here that, in some cases, the social ties of mutual obligation that facilitate the flow of material in the first place also install nonflows to preserve the sense of trust needed to collaborate. If the genetic biobank compiled in Copenhagen is to become a “platform” for *future* research,⁴⁰ we have shown here how it is also a materialization of *past* collaborations and acts of care heavily dependent on social skills.

We believe that it is important to acknowledge the practical kind of work that we have called “ethics work.” By staying aware of the tacit ethics work, we might better acknowledge all the “emotional labour”⁴¹ researchers need to do to make genetic material flow. Data sharing among participants and researchers does not just happen; it is made to happen. A lot of care is needed for the scientific enterprise to connect to the people in whom the samples and health data originate, and to create relations of trust between researchers.⁴² The making of Big Data depends on a form of detailed attention to local concerns that grand harmonized ethics policies do not ensure.

The accumulation of samples and data cannot be limited to the compiled material, knowledge, or machinery typically associated with “data sharing,” and the social skills on which flows of genetic material depend should not pass unnoticed. Such skills reflect the everyday ethos of work for researchers rather than the ethics policies intended to guide the international networks

³⁹ General Secretariat of the Council, “Council Conclusions on the Transition towards an Open Science System. Brussels, Council Of The European Union” (2016 May 27, REF 2596/16).

⁴⁰ See n 1.

⁴¹ Arlie Russell Hochschild, “Emotion Work, Feeling Rules, and Social Structure” (1979) 85 *American Journal of Sociology*, 551–75.

⁴² Carrie Friese, “Realizing Potential in Translational Medicine” (2013) 54(S7) *Current Anthropology*, 129–38; Annemarie Mol, *The Logic of Care: Health and the Problem of Patient Choice* (Routledge 2008).

through which they procure samples. The lab in Copenhagen has developed ways of keeping records of social concerns and modes of collaborating with international partners. It represents a form of ethics work that cares for detail. The emphasis in the lab's global collaborations is on the establishment of *enduring* engagements with partners and research participants. This implies returning information to those who want it, inviting international partners to work in the lab, and (in the process) creating social ties extending beyond plain data sharing. The tacit aspects of research collaborations remain largely unfunded, however, and this type of work does not provide much academic credit. If we begin acknowledging this important ethics work, we might also consider including it in new applications, discussing it in guidelines, and building future policies on experiences with this very practice-oriented form of ethics. This implies that funding agencies and research regulators should not impose demands of data sharing without paying attention to, and ensuring the necessary funding for, the work going into ensuring sustainable social relations among research participants and research partners. As such, the building of research infrastructures should focus more on fostering and nurturing social relations among actors.

The tacit forms of ethics work just described deserve dissemination to other fields seeking to implement genetic counseling based on emerging "Big Data" tools; for example, in attempts to introduce whole genome sequencing and exome sequencing into everyday clinical care.⁴³ The first step is to make it visible—and to make ethics debates focus on the mundane practices of care just described—with the same level of attention as has been paid to the abstract ethics principles guiding existing policy work. Only when we discuss what has already been done may we aspire to make data sharing policies and research ethics rules relevant for the science practices they are supposed to guide. In relation to the mobility turn and an increased interest in Big Data,⁴⁴ close scrutiny of actual flows will not only contribute to the study of data sharing policies; it will also form part of a larger corpus of scholarship exploring the preconditions for, and implications of, an ever more data-intensive research paradigm.

⁴³ Stefan Timmermans, "Trust in Standards: Transitioning Clinical Exome Sequencing from Bench to Bedside" (2015) 45(1) *Social Studies of Science*, 77–99.

⁴⁴ Sheller and Urry, n 21.